AD_____

Award Number: DAMD17-02-1-0373

TITLE: Modular Machine Learning Methods for Computer-Aided Diagnosis of Breast Cancer

PRINCIPAL INVESTIGATOR: Jonathan Jesneck
Dr. Joseph Lo, Ph.D.

CONTRACTING ORGANIZATION: Duke University
Durham, NC 27710

REPORT DATE: May 2004

TYPE OF REPORT: Annual Summary

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	May 2004	Annual Summary (1 May 2003 - 30 Apr 2004)

4. TITLE AND SUBTITLE
Modular Machine Learning Methods for Computer-Aided Diagnosis of Breast Cancer

5. FUNDING NUMBERS
DAMD17-02-1-0373

6. AUTHOR(S)
Jonathan Jesneck
Dr. Joseph Lo, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Duke University
Durham, NC 27710

E-Mail: jonathan.jesneck@duke.edu

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 Words)

The purpose of this study was four fold: 1) Identify subsets of the training data of breast cancer features using both a priori information and unsupervised learning methods. 2) Build local models for breast cancer prediction for each subset of the training data using supervised learning methods. Evaluate the performance of the local models on the training data relative to a single, global, supervised learning model and to current clinical practice. 3) Combine the local models to form a global, modular model and evaluate the performance of the modular model on the evaluation data set relative to a single, global, supervised learning model and to current clinical practice. 4) Develop an ensemble classifier combining three sources of data (image processing, radiologist-extracted mammographic findings, and patient history) for the task of computer-aided diagnosis of breast microcalcification clusters. We developed the world's largest database of over 4400 cases containing radiologist-extracted mammographic findings, patient age, and biopsy outcome, and we used this data to develop modular, global CAD models using different machine learning algorithms applied to the entire database.

14. SUBJECT TERMS
computer-aided diagnosis, breast cancer, BI-RADS, image processing, emsemble classifier

15. NUMBER OF PAGES
7

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

Table of Contents

Cover.....1

SF 298.....2

Introduction.....4

Body.....4

Key Research Accomplishments.....6

Reportable Outcomes.....7

Conclusions.....6

References.....7

Appendices.....N/A

INTRODUCTION

This study investigated modular and ensemble systems of machine learning methods for computer-aided diagnosis (CAD) of breast cancer to reduce the number of benign biopsies. While mammography is valuable for early detection of breast cancer, it has a high false-positive rate. A CAD system for identifying very likely benign lesions as candidates for follow-up instead of biopsy could spare women discomfort, anxiety, and expense and potentially improve the cost-effectiveness of mammographic screening programs.

This predoctoral fellowship covers two different students both mentored by Joseph Lo. It was originally awarded to Mia Markey, who graduated in 2002 from Duke University with her Ph.D. *Modular Machine Learning Methods for Computer-Aided Diagnosis of Breast Cancer*. The original aims were concluded as part of that dissertation research. As noted in last year's report, the Army authorized the transfer of the remaining fellowship to Jonathan Jesneck. We propose new aims 4 and 5 based on the success as well as difficulties discovered previously. Consistent with those aims, Mr. Jesneck has developed ensemble classifiers for the task of computer-aided diagnosis of breast microcalcification clusters, which are very challenging to characterize for radiologists and computer models alike. The rationale and progress for these aims is summarized in the report below.

BODY

The data consisted of mammographic features extracted by automated image processing algorithms. The same cases were used as described in last year's report.

Task 1. Identify subsets of the training data using both a priori information and unsupervised learning methods.

The database of digitized mammograms has already been created and analyzed, as described in last year's report. The most important grouping was mass vs. calcification lesions. In particular, both radiologists and computer models performed far worse when attempting to characterize the calcification lesions, which motivated the current emphasis on these types of lesions. This aim is now concluded.

Task 2. Build local models for breast cancer prediction for each subset of the training data using supervised learning methods. Evaluate the performance of the local models on the training data relative to a single, global, supervised learning model and to current clinical practice.

This task has already been completed and resulted in a publication, as described in last year's report. With regards to the challenging calcification cases, no local model was able to outperform the simple, single global model.

Task 3. Combine the local models to form a global, modular model. Evaluate the performance of the modular model on the evaluation data set relative to a single, global, supervised learning model and to current clinical practice.

This task has also been completed and published, as described in last year's report. The combination of modular models did not outperform the simpler, single global model. This negative result was attributed in part to the weaker performance over the challenging calcification cases. This again motivated the current work.

Task 4: For most challenging subset of the data, the microcalcification lesions, extract image-based features using a fully automated CAD scheme.

This task involved two major efforts. First, whereas all previous aims focused on radiologist-interpreted findings, here we extracted features from the digitized mammograms using a computer-aided detection (CAD) algorithm. These new features were incorporated into our first ensemble models, each based upon a different subset of features such as radiologist-interpreted, image processing, and patient history. Initial results were presented and published, as described in last year's report. A key limitation of that previous work was that the microcalcification detection algorithm developed by Gavrielides, *et al.* [1,2], missed up to 20% of clusters in its initial detection stage and could therefore never include them in its characterization stage. The bulk of effort this year has been devoted to developing a new algorithm to improve the sensitivity of the detection of microcalcifications. Rather than segmenting the calcifications based on the shape of the histogram of a local window of the preprocessed image, our new algorithm detected calcifications using the signal difference to noise ratio of the local window. The signal content of each image was first enhanced by subtracting away a median filter blurred version of itself. To keep only the high spatial frequency content of each window, all pixel values below three times the standard deviation of the background median-filtered window pixels were set to zero. The remaining pixels were considered to be possible calcifications, although they also included false positives such as due to ligament or duct crossings. This method was designed to be more sensitive but less specific than the previous calcification detection algorithm. With no false positive reduction stage, the algorithm performs at 96% sensitivity and 53.85 false positive clusters per image. While there are many false clusters per image, the sensitivity is high enough to allow characterization of microcalcification lesions at a clinically acceptable sensitivity. In on-going work, we plan to reduce the number of false positive detections with a two-pronged approach. First, we will optimize the initial thresholding to improve specificity of the signal-enhanced image prior to clustering. Second, we will adapt our existing false positive reduction algorithms to the clustered results, with the goal to eliminate the false positive clusters while keeping the true positive clusters. As a very preliminary attempt at false positive cluster reduction, linear discriminant analysis (LDA) models were used on the image-based features to determine their ability to distinguish between malignant clusters vs. normal breast tissue corresponding to computer false positive detection locations. The morphological features alone yielded an AUC of 0.77 ± 0.1, while the LDA using both morphological features and Haralick features with an angle of 45 degrees and a distance of 1 pixel performed with an AUC of 0.79 ± 0.02. Unfortunately, these results were not adequate for either detection or characterization of the lesions. The features for these models are described below.

Morphological features

The morphological descriptors of the individual microcalcifications were area, mean density, eccentricity, moment ratio, axis ratio, and the number of microcalcifications in the cluster. These features have been shown to aid in CAD schemes for microcalcification clusters [3]. The cluster morphological features were summaries of the individual microcalcification morphological feature values. These summary statistics were the maximum, average, standard deviation, and coefficient of

variation, which is the ratio of the standard deviation to the average. This resulted in a total of 20 morphological cluster features.

Texture features

The spatial gray-level dependence (SGLD) matrix is the joint probability of the occurrence of gray levels for pixel pairs which are separated by a particular distance and at a particular angle [4]. The 13 SGLD or co-occurrence matrix features are correlation, entropy, energy (angular second moment), inertia, inverse difference moment, sum average, sum entropy, sum variance, difference average, difference entropy, difference variance, information measure of correlation 1, and information measure of correlation 2. The SGLD matrices were calculated over the bounding boxes of the detected microcalcification clusters. Angles of 0 degrees and 45 degrees and distances of 1 pixel and 10 pixels were considered. Additional angle and distance combinations are currently being investigated.

Task 5: Develop ensemble models for predicting benign vs. malignant calcification clusters.

As discussed in last year's report, modular systems tended to match but not exceed the performance of a classic feed-forward, back-propagation artificial neural network. As a result of that last discovery, we are focusing efforts now on developing ensemble models for calcifications. We are currently investigating various models and automated feature selection methods to reduce the false positive cluster rate. The next steps will involve using the models to characterize the clusters as either benign or malignant and adding to the new models the BI-RADS features and patient age, as discussed in last year's report.

KEY RESEARCH ACCOMPLISHMENTS

- Developed a new algorithm for microcalcification detection with high sensitivity
- Added morphological features for microcalcifications and microcalcification clusters
- Added texture features for microcalcification clusters
- Developed a preliminary fully automated ensemble CAD system to detect microcalcification clusters

CONCLUSIONS

The current work focuses on a change in strategy towards image based feature information and particularly for microcalcification lesions, but the final goal remains the same, namely to develop models to predict the outcome of breast biopsy. The current results are very much a work in progress, with good sensitivity but still lacking in false positive reduction and also weak characterization performance. We hypothesized that the basic limitation was the excessive number of false positives, which provided poor input examples for the LDA models. We further hypothesized that we did not take advantage of the great variety of distances and angles which would result in many more texture features. By improving the front-end detection specificity and the back-end characterization using better texture features, and by merging in the additional contribution of radiologist-interpreted and patient age, we anticipate being able to improve greatly upon these results in the coming year.

REPORTABLE OUTCOMES

None.

REFRENCES

1. M Gavrielides, JY Lo, CE Floyd, Jr., Parameter optimization of a computer-aided diagnosis scheme for the segmentation of microcalcification clusters in mammograms, Med. Phys, 29(4):475-483, 2002.
2. MA Gavrielides, JY Lo, R Vargas-Voracek, CE Floyd, Jr., Segmentation of suspicious clustered microcalcifications in mammograms, Med. Phys., 27(1):13-22, 2000.
3. H Chan, B Sahiner, KL Lam, N Petrick, MA Helvie, MM Goodsitt, DD Adler, Computerized analysis of mammographic microcalcifications in morphological and texture feature spaces, Med. Phys. 25(10):2007-2019, 1998.
4. RM Haralick, K Shanmugam, I Dinstein. Textural features for image classification. IEEE Transactions on Systems Man and Cybernetics, SMC-3(6):610-621, 1973.